STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/553,386_
Source: _IFWP_
Date Processed by STIC: _1/11/07_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05): U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ERROR DETECTED	**SUGGESTED CORRECTION**	**SERIAL NUMBER:** _10/553 386_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length
The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino Numbering
The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII
The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length
Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0 "bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences (OLD RULES)
Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences (NEW RULES)
Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's (NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213> Response
Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence. (see item 11 below)

11 ____ Use of <220>
Sequence(s) ___ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section or use "chemically synthesized" as explanation. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32), also Sec. 1.823 of Sequence Rules

12 ____ PatentIn 2.0 "bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n/Xaa
"n" can only represent a single <u>nucleotide</u>; "Xaa" can only represent a single <u>amino acid</u>

IFWP

RAW SEQUENCE LISTING   DATE: 01/11/2007
PATENT APPLICATION: US/10/553,386   TIME: 09:47:27

Input Set : A:\PTO.RJ.txt
Output Set: N:\CRF4\01112007\J553386.raw

Does Not Comply
Corrected Diskette Needed

see item 4 on Erro Summary Sheet

```
      2 <110> APPLICANT: Hallenbeck, Paul
      3         Hampton, Garret
      4         Hay, Carl
      5         Huang, Ying
      6         Jakubczak, John
W-->  7 <120> TITLE OF INVENTION: TMPRSS2 REGULATORY SEQUENCES AND USES THEREOF
W-->  8 <130> FILE REFERENCE: 3802-166-27 NATL
C-->  9 <140> CURRENT APPLICATION NUMBER: US/10/553,386
C-->  9 <141> CURRENT FILING DATE: 2005-10-17
      9 <150> PRIOR APPLICATION NUMBER: 10/553,386
     11 <160> NUMBER OF SEQ ID NOS: 8
     12 <170> SOFTWARE: FastSEQ for Windows Version 4.0
```

delete - this is not a prior application number

ERRORED SEQUENCES

```
W--> 13 <210> SEQ ID NO: 1
     14 <211> LENGTH: 1811
     15 <212> TYPE: DNA
     16 <213> ORGANISM: H. sapiens
W--> 17 <400> SEQUENCE: 1
     18 agagcttccc aggaagggag cagtgagcca aggcagcctg ggatgggact gaatggggct 60
     19 tttttctgct tccacctcat tttaaagcaa atcatgttga tttgtatatt atgcagggga 120
     20 ggagaaaaca tgcttcccct ctacccttct agttcgttca gctagtctac aaattaaatt 180
     21 ggcatcaaat agatgaacag gagaaaagct gttttaatgt atgtactcac agatgggaat 240
     22 cccacaagaa tatgagactt aaagaacagg ccaggtgagt gaggggtcca gtgcgggggc 300
     23 tcacacctgt aatcccagca ctttgggagg ccaaggcagg tggatcgctt gagcccagat 360
     24 tggagaccag tctgggtaac atagggagac ccccatctct acaaaaaatt aaattagctg 420
     25 ggcgtggtgg cacacgcctc tagtcccagc tactcgggag gctgaggtgg gaggattact 480
     26 tgtgccctgg gaggttgaag ctgccatgag ccatgattga gccactgcac tccaacctgg 540
     27 gcaacagagt gagacccctt ctgggggaaa aaaaaacaca cgaaaaaaaa aaggtgcagc 600
     28 agcccgatga ttgaggctta tctgtcactc tgagtgacag aaagaaatgg gggtttgagg 660
     29 cttctgggga gcggtggagg agtgagggga gcgtgaggag aggaggtgtc tggtgaacgc 720
     30 aggttgccgt gtgaggcaga taaaagtttc ccaggtgata aaagttgtcc gggaacagct 780
     31 ctcttcctgg tacagatctg ctgactaaca aacatttcct ttataggtgc aaatttcctt 840
     32 tacaaaaggg cattttctca gaggtactct ggtgtctgca gttcctcaac ataaccagtt 900
     33 ccaaatcatc aatgtgccaa agaggactat gttggggtag cagattctgg tctcctccag 960
     34 tcctacttgg ggtgatgaat tctggtctac ggtcctatta aattctggtg aattctgagt 1020
     35 ccccacaatt gcaaacatta gaaagaacct ctcaagtgcc cgggaacagc cacgtcttcc 1080
E--> 36 tggctgaggt gtgtcgacc acttcctcac tcccgccctg gccggtggtg ccgagagacc 1140
     37 tgggaccatc cggggagcc ctttccaccg gacgctggtg ggggccaaga aatgccagcc 1200
     38 taggcggact ggggagggtc ttgggcgtcc ggcgctgtgt ccccgccact cgtgcttggg 1260
E--> 39 ccagcagtcc ccaaggccta ctcctgggtc cttgcccaga ggctacagtg ggttccccgg 1320
```

invalid nucleic acid designator

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/553,386

DATE: 01/11/2007
TIME: 09:47:27

Input Set : A:\PTO.RJ.txt
Output Set: N:\CRF4\01112007\J553386.raw

```
      40 aggccaagac ggggccggcc gcctacagga gctcgtgagg tagcagctcc gggggctcac 1380
      41 ccaggactcc aggagcgctc cccagaatcc ccttccttaa cccaaactcg agccctcggg 1440        )invalid
E--> 42 cagcgctgcg ccgcggaccg gagaggggca ggttggccgc tgtggccggg cccgggaagc 1500
      43 gccccagagt ccccttatggg tccctccgca gccggggttg agccaggcag ggaacccgtc 1560
      44 cggacttccc ttgggaaacg cctcctcccg ccgcccccgc ccccgcccgc ccaggtgac 1620
      45 ccgcgacccg cttgggggtg tcgccctgga ccctgggaca ccgcctcctg agattaaagc 1680
      46 gagagccagg gcgggccggg ccgagtaggc gcgagctaag caggaggcgg aggcggaggc 1740
      47 ggagggcgag gggcggggag cgccgcctgg agcgcggcag gtgagcggcg ccggtaccag 1800
      48 ggtcccggct c 1811
      49 <210> SEQ ID NO: 2
      50 <211> LENGTH: 239
      51 <212> TYPE: DNA
      52 <213> ORGANISM: H. sapiens
      56 <400> SEQUENCE: 2          invalid
      57 cgtggcggag ggactgggga cccgggcacc cgtcctgccc cttcaccttc cagctccgcc 60
      58 tcctccgcgc ggaccccgcc ccgtcccgac ccctcccggg tccccggccc agccccctcc 120
      59 gggccctccc agcccctccc cttcctttcc gcggccccgc cctctcctcg cggcgcgagt 180
E--> 60 ttcaggcagc gctgcgtcct gctgcgcacg tgggaagccc tggcccggc caccccgc 239
      95 <210> SEQ ID NO: 8
      96 <211> LENGTH: 270
      97 <212> TYPE: DNA
      98 <213> ORGANISM: H. sapiens
W--> 99 <400> SEQUENCE: 8
     100 tggtaccatc cggacaaagc ctgcgcgcgc cccgcccccgc cattggccgt accgccccgc 60
     101 gccgccgccc catcccgccc ctcgccgccg ggtccggcgc gttaaagcca ataggaaccg 120
     102 ccgccgttgt tcccgtcacg gccgggggcag ccaattgtgg cggcgctcgg cggctcgtgg 180
     103 ctctttcgcg gcaaaaagga tttggcgcgt aaaagtggcc gggactttgc aggcagcggc 240
     104 ggccgggggc ggagcgggat cgagccctcg 270
E--> 110 ganF1\359709.1
E--> 112   - 1 -          delete
```

VERIFICATION SUMMARY   DATE: 01/11/2007
PATENT APPLICATION: **US/10/553,386**   TIME: 09:47:28

      Input Set : **A:\PTO.RJ.txt**
      Output Set: **N:\CRF4\01112007\J553386.raw**

```
L:7 M:283 W: Missing Blank Line separator, <120> field identifier
L:8 M:283 W: Missing Blank Line separator, <130> field identifier
L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:13 M:283 W: Missing Blank Line separator, <210> field identifier
L:17 M:283 W: Missing Blank Line separator, <400> field identifier
L:36 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:39 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:42 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:60 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:65 M:283 W: Missing Blank Line separator, <400> field identifier
L:75 M:283 W: Missing Blank Line separator, <400> field identifier
L:81 M:283 W: Missing Blank Line separator, <400> field identifier
L:87 M:283 W: Missing Blank Line separator, <400> field identifier
L:93 M:283 W: Missing Blank Line separator, <400> field identifier
L:99 M:283 W: Missing Blank Line separator, <400> field identifier
L:110 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:8
L:110 M:254 E: No. of Bases conflict, LENGTH:Input:1 Counted:276 SEQ:8
L:110 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:10
L:110 M:112 C: (48) String data converted to lower case,
M:254 Repeated in SeqNo=8
L:112 M:252 E: No. of Seq. differs, <211> LENGTH:Input:270 Found:276 SEQ:8
```